Sinequa. Your Company. Augmented.

Accelerating the Life Sciences with Cognitive Search

Search has revolutionized the way we interact with the Internet. What used to be an exponentially expanding ocean of uncharted information has become a valuable resource we can plot a course through – so long as we have a basic idea of what we're looking for. In our attempts to sail this endless, publicly-accessible sea, we now perform 99,000 Google searches per second. But there are other large sources of digital data that remain opaque. Businesses often have untapped sources of information spread throughout their internal tools, databases, and other repositories that could help drive their success, but they don't have a way to find what they need within them. This is especially true of knowledge-focused industries such as the Life Sciences.

In 2020, every person in the world created an average of 1.7MB of data per second, and 90% of the world's data was generated in the last two years. Every day, 306 billion emails are sent, and that will increase to 361 billion by 2024. The volume of data we produce is growing exponentially, creating a huge wealth of information that can be mined for insights. Social media companies, for example, have controversially turned the data their users produce into a resource to optimize their advertising revenue.

But as much as there is a copious amount of data on the public Internet, even more resides behind the firewalls of private companies. Whereas the public Internet is now extremely well indexed – Dark Web excluded – corporate data is often unsearchable and hidden from view. Not only is it protected from external access, as it should be, but this data is frequently opaque to internal employees as well. This is where Sinequa's Intelligent Enterprise Search can pay dividends, turning internal corporate data into a valuable resource for business. For Life Sciences companies such as the pharmaceutical industry, the rewards generated can be immense.

Empowering drug discovery with cognitive search

The value built by a Life Sciences organization comes from the discoveries it makes. These benefit mankind as well as having considerable potential for financial rewards. One of the main activities involved will be discovering new drugs. Activities like drug discovery are about exploring possibilities. They don't merely involve matching a compound with a given range of symptoms. The Life Sciences move forward via a creative process where researchers need to assess valid molecule candidates for potential new capabilities, which can then be tested further. Side effects, costs, and availability must also be taken into account.

This process involves a lot of lab experiments. These are the lifeblood of generating any scientific knowledge. A key part of experimentation is reproducibility, where a given set of conditions is shown to create results that can be repeated. This enables certainty about the ability of a drug to provide reliable therapeutic benefits. The records from these lab tests need to be clear and easy to find, otherwise researchers won't know what experiments have been performed in the past and what happened. Each new experiment builds on the results of previous ones.

Research is to see what everybody else has seen, and think what nobody has thought

Albert Szent-Györgyi

The problem is that this data is held in many different repositories. These include Electronic Lab Notebooks (ELN), Lab Information Management Systems (LIMS), Chromatography Data Systems (CDS) and Laboratory Execution Systems (LES) to name a few. Not only are there many data silos, but the interfaces to find anything within each one can be poor. There will also be data in multiple repositories that relates to each other but isn't directly connected. Even a single repository may not make it so easy to find important information. For example, the data in ELNs can provide ongoing project knowledge as a drug discovery process progresses. An ELN captures detailed information about experiments as well as analysis and interpretation of results. This information can be in the form of text notes, tables of data, image attachments, and links to relevant data from previous years. A lot of useful insight could be available here, but its organization might not be consistent from one researcher to the next, or the contents may be incomplete.

This can cause issues with new and existing drug discovery research that can hinder and delay development. It could be hard to find data from an experiment that was performed some years ago. Or a researcher may want to replicate an earlier experiment but can't find the full details about it. If the colleague who performed the earlier experiment has left the organization, it may be difficult or impossible to track these details down. Similarly, the person who performed the experiment may not be around for other reasons – perhaps they are on leave at a crucial moment or on sabbatical.

A researcher may not even know if their organization (or another organization) has worked on a target drug compound before. This could lead to unnecessary duplication of testing or might even fail to reveal that this target has already been ruled out and research should look elsewhere. Drug discovery may have its foundation in lab experimentation, but as this is the most expensive and time-consuming part of the process, it's paramount that effort is focused on the most relevant candidate molecules. It's therefore crucial to have a clear picture of past experiments and their results. A comprehensive search facility that spans all the relevant repositories will deliver a more effective process. This is what Sinequa's Intelligent Enterprise Search provides for Life Sciences companies with its cognitive approach.

Enhancing search serendipity

Drug discovery experiments in labs are supported by a development setup for initial intelligence gathering that allows the focus to be on applicable molecules. The process looks at internal and external data sources for efficacy, selectivity, safety, academic context, and what the competition is up to via external research. Sinequa's Intelligent Enterprise Search facilitates this by bringing all the relevant sources into one interface. Some of these sources will be structured, so will already have organizational categories applied that can be utilized.

But 80% of world data is unstructured, and this is just as true of the Life Sciences industry as it is in other domains. Sinequa Intelligent Search's ability to ingest data from a huge variety of unstructured sources is a key factor in the value it brings. Sinequa can universally connect to any data source required by a customer and supports over 350 file formats. This means a plethora of data can be brought into the search solution, transforming it into relevant information in context. The application of Sinequa's sophisticated Natural Language Processing (NLP) and AI/ML-powered semantic analysis make it possible to extract the most meaning from unstructured sources.

Not finding the right information fast enough can cause costly delays in drug discovery. According to research by the London School of Economics, the average cost of bringing a new drug to market is $1.3 billion. Spending too long finding the right information regarding which molecules to test can significantly increase this figure. By implementing Sinequa's Intelligent Enterprise Search, a pharmaceutical company can speed the process up by making results much more relevant and complete. This is built on the platform's solid basis of being able to connect with every possible data source that could be included.

The Sinequa platform is configured for the drug industry by enabling a market-wide search during initial setup covering both internal and external sources. The data from these sources is enriched and structured using NLP, semantic analysis, text mining, classification, and other methods. This ensures all information about promising compounds is included in a search, and all relevant data can be reviewed before experimentation begins.

The application framework and component library supplied by Sinequa make it easy for customers to tailor their own end-user applications. They can also use the platform's REST API to integrate data derived using Sinquea's technology into their other software. Statistical and neural relevance engines enable users to curate their own information journeys, focusing on the data types that are most relevant to their business' priorities.

A unique capability Sinequa Intelligent Enterprise Search provides is uncovering the experts within a domain. Alongside relevant data links, results include the most frequent human sources of insightful research. Although drug discovery is a scientific process based on facts, knowing who the experts are in an area enables researchers to enlist them for additional insights. These human conversations can improve understanding more quickly than reading research documents on their own. The same facility can be used for building research teams for a new project, by pinpointing the experts in a specialized area.

The security of data is also becoming increasingly important, as privacy regulations such as Europe's GDPR and California's CCPA have come into force. Sinequa's search technology is designed to respect native system security but also leaves data intact in its original location, rather than making a copy.

With Sinequa, we are building a powerful next-generation search platform that is simple and intuitive enough for our R&D scientists to use easily and be alerted to new information anywhere, anytime

Albert Szent-Györgyi

In any business related to healthcare that handles patient data, this is a crucial capability. It's also an essential feature in a competitive business such as pharmaceuticals, where discovering a cure first and patenting the drug compounds involved ahead of the competition provide the route to a corporation's success.

Tried and trusted technology

Sinequa's technology is already providing value to the Life Sciences, with 50% of the world's largest pharmaceutical companies already using Sinequa Intelligent Search, including Pfizer, AstraZeneca, GSK, Novartis, and Bristol Myers Squibb. These organizations are seeing the impact of being able to seamlessly search and navigate a multitude of data sources without disrupting or compromising the existing security access rights of the source systems.

The benefits span all the core domains related to the Life Sciences industry. Drug discovery teams gain an easy, unified way to call upon the collective knowledge of their co-workers, their company, and the wider scientific community. They can identify relevant and useful innovations that have occurred both within their own organization and in the external research world. They can avoid repeating mistakes or chasing a dead end, while pinpointing the experts in their field based on contribution. This can accelerate the journey to life-saving therapies.

Clinical statisticians can gain a holistic view across all experimental trial data and reports. They can employ this to find the previously unseen connections that uncover the right subjects or modifiers, and related side effects. They can locate subjects with certain diseases through their symptoms and then narrow these down by applying a plethora of filter criteria. They can perform these searches across drugs and studies.

Regulatory teams can gain insight into previous responses to queries submitted to global agencies such as the US Food and Drug Administration, European Medicines Agency, or the Pharmaceuticals and Medical Devices Agency. This can facilitate New Drug Applications by preventing the need to collect information from scratch for every response. Sinequa enables full-text searches on prior regulatory submissions and related documents from one interface. Regulatory query responses are unified across time and around the globe, speeding up the process and strengthening consistency.

Medical Information groups can harness the benefits of Sinequa Intelligent Enterprise Search when delivering scientific and clinical knowledge about a company's products. Gaining a uniform view of field data, which was traditionally distributed across discrete silos, saves time and enables employees to speak with a unified corporate voice. All employees will be working from a single portal to reduce the risk of conflicts with regulatory bodies.

Even at the manufacturing level, pharmaceutical lab personnel can enjoy a single mobile-enabled digital point of entry to knowledge about their equipment, powered by Sinequa. This is a massive improvement over paper-based systems. Facility operators, technicians, engineers, and process specialists gain easier access to information. The relevant documentation and standard operating procedures are digitized and indexed, so they can be made available through a portal accessible via a mobile device. User can search by keyword or QR code for relevant documentation using their smartphone or tablet.

These wide-ranging capabilities have prompted Gartner to name Sinequa a leader for Insight Engines in its 2015, 2017, 2018, 2019, 2021, and 2022 Magic Quadrants, while Forrester Wave came to the same conclusion in its Cognitive Search category in 2015, 2017, 2019 and 2021. Existing customers have praised the company's technology as well.

Nick Brown, Head of AI & Data Science, CTO Office,

AstraZeneca, said: "With Sinequa, we are building a powerful next-generation search platform that is simple and intuitive enough for our R&D scientists to use easily and be alerted to new information anywhere, anytime." Albert Bourla, CEO, Pfizer, said: "We are initiating an enterprise-wide digital effort to speed up drug development, enhance patient and physician experiences and access and leverage technology and robotics to simplify and automate our processes." Sinequa is a core element of this Pfizer strategy.

Sinequa Intelligent Enterprise Search is the only product on the market today that equips pharmaceutical organizations with the means to accelerate innovation, efficiency, and agility by connecting to all sources and synthesizing all research data in all formats and most languages into accessible and useful information, securely and at scale. Sinequa's customers are using it to keep up to date with the scientific knowledge around their drug discovery processes, answer hard questions quickly during drug development, and accelerate the design and management of clinical trials.
The benefits of adopting a Sinequa solution in Life Sciences drug discovery are readily quantifiable. Top of the list is the decreased time required to bring drugs to market. Intelligent Enterprise Search can also increase revenue by speeding up the delivery of marketing materials to local markets. It can increase potential income by identifying research opportunities within old data. Another boost to revenue can come from identifying the websites of fraudulent companies so these can be pursued and terminated. The more comprehensive search can decrease regulatory and privacy compliance risk by providing a wider search for areas of vulnerability, which can potentially save millions. Overall, project teams can be more productive, with faster composition of expert groups.
Sinequa Intelligent Enterprise Search isn't just a useful extra service for a Life Sciences company to have. The solution has real tangible benefits both economically and for how quickly it can deliver effective cures for human ailments. It has already found favor with some of the biggest names in the pharmaceutical industry, many of which have made Sinequa a key element

in the information technology strategy. These customers have continued to use Sinequa technology since first adoption and have universally increased their investment. They are realizing quantifiable results that deliver faster times to market with new drugs. Ultimately, Sinequa Intelligent Enterprise Search provides the technology Life Sciences companies need to make the most effective use of their existing knowledge capital.